# F-structure and s-structure of Urdu complex predicates

Joseph Lovestrand
SOAS University of London

Proceedings of the LFG'20 Conference

On-Line

Miriam Butt, Ida Toivonen (Editors)

2020

CSLI Publications

pages 232–249

http://csli-publications.stanford.edu/LFG/2020

Keywords: Urdu, complex predicates, glue semantics, argument structure, a-structure, s-structure, semantic forms

Lovestrand, Joseph. 2020. F-structure and s-structure of Urdu complex predicates. In Butt, Miriam, & Toivonen, Ida (Eds.), *Proceedings of the LFG'20 Conference, On-Line*, 232–249. Stanford, CA: CSLI Publications.

**Abstract**

This paper proposes a solution to the problem of constituent-sensitive interpretation of semantic scope in Urdu complex predicates in the analysis of Lowe (2015). In this novel approach, the PRED value of the main verb in an Urdu complex predicate appears in its own f-structure with no grammatical functions. The lexical entry of the light verb constructs a complex s-structure which links the single set of grammatical functions with semantic arguments in a connected s-structure. This approach avoids the formal complications of the hybrid f-structure proposal of Andrews (2018). Resolving the semantic scope problem provides support for the idea that complex predicates are formally unproblematic in LFG, as long as argument selection is constrained by glue semantics (instead of general Completeness and Coherence constraints). This approach to complex predicates is a possible alternative to the approach of Butt (1995) which postulates the formal complications of a-structure and unifiable semantic forms.

# 1 Introduction

Asudeh and Giorgolo (2012) posit a version of the LFG architecture that builds on the idea that glue semantic constraints can effectively replace Completeness and Coherence (Dalrymple et al. 1993, Kuhn 2001, Andrews 2008). They argue that, under this assumption, there is no need for a level of a-structure (distinct from s-structure) located between c-structure and f-structure, as in the architecture posited by Butt, Dalrymple and Frank (1997). Asudeh and Giorgolo (2012: 69-72) further point out that a level of a-structure separate from s-structure significantly complicates the architecture, in particular by requiring two separate links between a-structure and s-structure, one via f-structure and one direct (if no overt argument is expressed in f-structure; see Findlay (2016: 303-309) for further discussion). The basic architecture proposed by Asudeh and Giorgolo (2012) is illustrated in Figure 1.

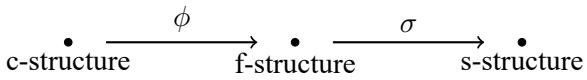



Figure 1: LFG architecture (Asudeh and Giorgolo 2012)

Several researchers have worked on developing different aspects of this architecture: Asudeh and Giorgolo (2012) on optional arguments, Asudeh et al. (2014) on valency and derived arguments, Findlay (2016, 2020) on function-argument linking, Przepiórkowski (2017) on the argument-adjunct distinction, as well as

[†]Thanks to the participants of the LFG20 conference for their constructive feedback, in particular: Ash Asudeh, Avery Andrews, Miriam Butt and John Lowe; as well as to several anonymous reviewers of this paper and the original conference abstract. Naturally, any remaining faults are my own.

Lovestrand (2018) on argument sharing in serial verb constructions. Of particular interest in this regard is Lowe's (2015) analysis of Urdu complex predicates. Data from Urdu complex predicates are closely tied to the original a-structure architecture proposal. If complex predicates cannot be analyzed without a-structure, then the complexity introduced into the architecture by a-structure might still be exactly the level of complexity needed. The simpler architecture of Asudeh and Giorgolo (2012) is only as good as the most complex type of construction it can represent.

## 2 Urdu complex predicates

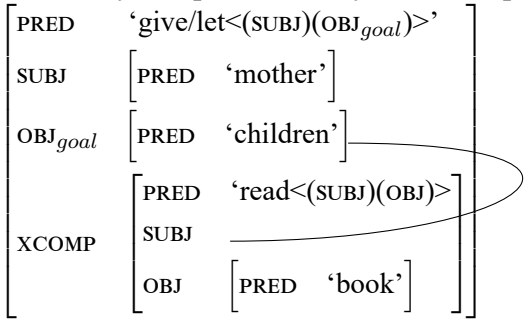
Complex predicates in Urdu/Hindi have been described and analyzed in extensive detail by Butt and colleagues (e.g. Butt 1995, 1997, Butt, King and Maxwell III 2003, Butt, King and Ramchand 2010, Butt and Ramchand 2005, Butt 2014). Butt's earlier work primarily examines two types of complex predicates:[1] Aspectual complex predicates which have an effect on the lexical semantics of the main verb without changing the valency, and permissive complex predicates which increase the valency of the main verb. Only the latter will be discussed here. In a permissive complex predicate, the light verb *de* 'give' increases the overall valency of the predicate by adding a "permitter" to the argument structure. In example (1), the main verb is *paṛʰ-ne* 'read'. Its agent is *baccõ* 'children', and its patient is *kıtab-ẽ* 'books'. The presence of the light verb *dĩ* 'give' coincides with an additional argument, *mã=ne* 'mother', understood to be the permitter, allowing the child to read the book.

(1) mã=ne bɑccõ=ko kıtab-ẽ pɑṛʰ-ne
mother(F)=ERG child(M).PL.OBL=DAT book(F).-PL.NOM read-INF.OBL
dĩ
give.PRF.F.PL
Mother let (the) children read (the) books. (Butt 2014: 2)

In example (1), grammatical relations correspond to case marking in a straightforward manner. The permitter is a subject, as indicated by its ergative case. The patient of the main verb, in nominative case, is an object, and the agent of the main verb, in dative case, can be considered a secondary object. The semantics of the Urdu permissive intuitively suggest (at least from an anglocentric perspective) that the patient is the grammatical object of the main verb, and that while the agent may be a secondary object in relation to the light verb, it functionally could also be a subject of the main verb. This "control" analysis is illustrated in example (2). However, the Urdu data contradicts this intuition.

[1] More detailed analyses of Urdu causatives are included in later work (e.g. Butt and King 2006, Butt et al. 2010, Butt 2014).

(2) **Intuitive yet impossible analysis of example (1)**




Despite the fact that the secondary object (OBJ$_{goal}$) is the semantic agent of the main verb, there is no evidence that it functions as a subject in any sense (Butt 2014: 10-13,18-19), so the control analysis in example (2) is unsupported by the evidence. Furthermore, the agreement patterns in example (1) indicate that the object (plural feminine) is the object of the light verb, despite being the patient of the main verb (Butt 2014: 13-15). An analysis of Urdu permissive constructions must account for the fact that there is just a single set of grammatical functions linked to a complex argument structure.

This configuration of semantic arguments and grammatical functions is particular puzzling for any theory of argument realization. It is typically assumed that a predicate determines its arguments (e.g. agent and patient) and that there is a standard process for determining (or constraining) which grammatical function each argument should have (i.e. Lexical Mapping Theory; Bresnan and Kanerva 1989). For a verb like 'read', the expected standard alignment is for the agent to be the subject and the patient to be the object. The light verb in the permissive construction does two things: it introduces a new argument that is not part of the lexical semantics of the main verb, and its presence coincides with a different linking: the agent of the main verb is no longer expressed as subject.

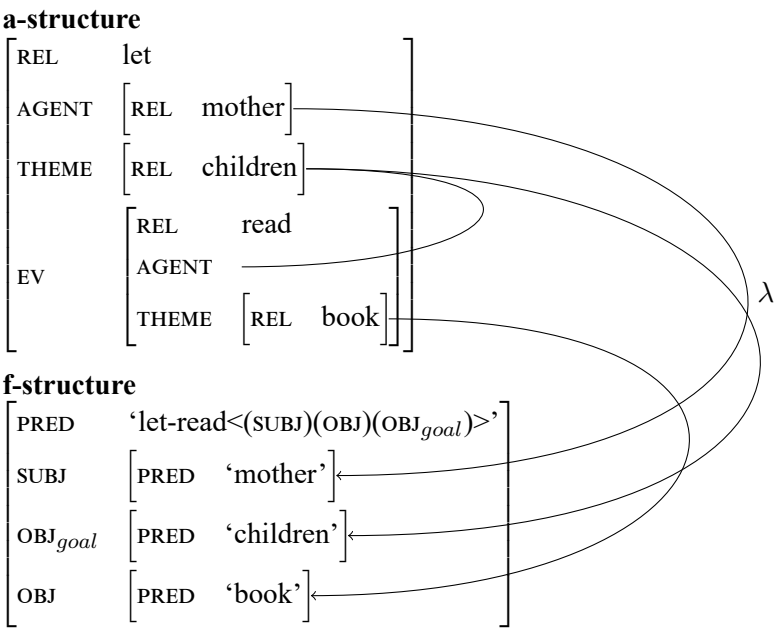
# 3 Butt's LFG analysis

Butt's (1995) analysis of Urdu complex predicates follows the traditional LFG assumption that arguments are licensed by the semantic form (the value of PRED) of a predicate in a subcategorization frame. The generalized constraints of Completeness and Coherence restrict the grammatical functions that can appear in the f-structure of that PRED to all and only the grammatical functions appearing in the subcategorization frame of the semantic form (Bresnan and Kaplan 1982). The implication is that the light verb in a valency-increasing complex predicate must have a PRED attribute. If both the light verb and the main verb have a PRED attribute, each with its own subcategorization frame, it is unclear how Completeness and Coherence can be satisfied if there is only a single set of grammatical functions in f-structure.

Butt proposes that complex predicates should be treated as a case of two PRED

values unifying in the syntax, modifying the original position of Bresnan and Kaplan (1982) in which semantic forms are considered non-unifiable. Assuming that at least some semantic forms can be unified, Butt develops a model in which the main verb's lexical entry retains its standard subcategorization frame, for example, 'read<(SUBJ)(OBJ)>', but this is altered when it interacts with a light verb in the syntax so that the agent is no longer linked to the SUBJ, but to the $\text{OBJ}_{goal}$. Butt (1995: 145) calls this "argument fusion" or "argument merger" (Butt 2014: 20).[2] In argument fusion, the lexical semantic structure of the main verb is embedded in the argument structure of the light verb, and at least one of the arguments of the embedded main verb is co-indexed with an argument of the higher verb in the resulting semantic structure. The result is an f-structure of a complex predicate with a single PRED attribute whose semantic form is composed from the two predicates, as shown in example (3). This provides a natural account for why only a single set of grammatical functions are found in the construction, under the standard assumption that the presence of grammatical functions is licensed by the subcategorization frame of the semantic form.

(3) **Analysis of example (1) based on Butt (1995) and Butt et al. (1997: 12)**




There have been two basic critiques of argument fusion. The first critique is that it has not been satisfactorily shown how this approach can be fully integrated into the formal assumptions of the LFG architecture (Lowe 2015: 419-423). The second critique is that it is a syntactic process that modifies a lexical entry. "A fundamental assumption of early lexicalist syntax was the principle of Direct Syn-

[2]Note that Butt's work builds on original proposals by Alsina (1993, 1997) working on causative constructions in Romance and Bantu languages.

tactic Encoding, i.e. the principle that lexical properties such as argument structure should not be manipulable in the syntax. This plays out in LFG in the fact that, at least originally, semantic forms are not manipulable in the syntax" (Lowe 2015: 418). Ideally, an analysis of complex predicates should be fully implementable and should not require modifying semantic forms in the syntax.

# 4 Lowe's LFG+glue re-analysis

Lowe (2015), following a similar proposal by Dalrymple et al. (1993), proposes an analysis of Urdu complex predicates within the architecture proposed by Asudeh and Giorgolo (2012). The relationships between the semantic arguments and their grammatical functions are modeled through meaning constructors in the semantics, without reference to a-structure. Example (4) shows a standard glue meaning constructor for the verb 'read'. On the left side of the colon, the semantic representation represents the relationship between the two arguments of the predicate. On the right side of the colon, the glue expression requires two kinds of semantic arguments, here labeled ARG2 and ARG1, in order to produce the meaning of the predicate. The lexical entry of the verb also contains optional equations, as in example (5), allowing the ARG1 to link to either a SUBJ or OBL; and the ARG2 to link to either a OBJ or $\text{OBJ}_\theta$ (for further discussion, see Findlay 2016, 2020).

(4) $\lambda y.\lambda x.read(x, y)$ :
$(\uparrow_\sigma \ \text{ARG2}) \multimap (\uparrow_\sigma \ \text{ARG1}) \multimap (\uparrow_\sigma \ \text{EV}) \multimap \uparrow_\sigma$

(5) a. $((\uparrow \text{SUBJ})_\sigma = (\uparrow_\sigma \ \text{ARG1}))$
b. $((\uparrow \text{OBJ})_\sigma = (\uparrow_\sigma \ \text{ARG2}))$

In Lowe's analysis, the meaning constructor of the light verb (example (6)) increases the valency of the construction while simultaneously ensuring that each argument is associated with the correct part of the semantic interpretation. On the meaning side, the light verb consumes a single argument and its predicate, and returns a predicate requiring two arguments, effectively adding an ARG3 to the argument structure. On the meaning side, the variable *P* represents the meaning of the main verb, and its meaning is embedded as an argument of the light verb's permissive predicate. The first argument, ARG3 is associated with the permittee role and the agent role of the embedded predicate. The ARG1 is assigned the permitter role.

(6) $\lambda P.\lambda y.\lambda x.\lambda e.let(x, y, P(y, e))$ :
$[(\uparrow_\sigma \ \text{ARG1}) \multimap (\uparrow_\sigma \ \text{EV}) \multimap \uparrow_\sigma] \multimap$
$(\uparrow_\sigma \ \text{ARG3}) \multimap (\uparrow_\sigma \ \text{ARG1}) \multimap (\uparrow_\sigma \ \text{EV}) \multimap \uparrow_\sigma$

Since subcategorization is handled by glue semantics, it is divorced from the PRED attribute. Under these assumptions, there is no need to assume that the light verb has a PRED attribute, even though it licenses the permitter argument in the

semantics. Lowe (2015: 427) represents the contribution of the light verb to f-structure with a binary feature PERMISSIVE, as in example (7).

(7) **F-structure of example (1) based on Lowe (2015)**

$$\begin{bmatrix} \text{PRED} & \text{‘read’} \\ \text{PERMISIVE} & + \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{‘mother’} \end{bmatrix} \\ \text{OBJ}_{goal} & \begin{bmatrix} \text{PRED} & \text{‘children’} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{‘book’} \end{bmatrix} \end{bmatrix}$$

The s-structure resulting from this analysis is similar to the f-structure. It is a flat structure containing a single ARG1, a single ARG2 and a single ARG3. The ARG3 in s-structure is linked to the $\text{OBJ}_{goal}$ in f-structure, as expected in a straightforward linking of grammatical function and semantic arguments. One potential disadvantage to this approach is that the s-structure representation does not model the complex meaning of the predicate. The permissive semantics are only found in the glue semantics, and not modeled at any level of representation, including s-structure. Whether or not this is problematic depends on one's view of s-structure. The approach proposed in Section 6 leaves open the possibility of treating s-structure as a model of complex semantic meaning, especially complex meaning composed of more than one predicate.

# 5 Andrew's proposed solution to the scope problem

Lowe (2015: 453) identifies a shortcoming in a strictly glue-based analysis. The approach overgenerates possible interpretations of (very) complex predicates with more than one light verb. Example (8) is an Urdu construction with a causative complex predicate embedded in an aspectual complex predicate embedded in a permissive complex predicate. There are two valency-increasing predicates, but it is the verb *di-ya* that is higher in scope: 'allow to cause', not 'cause to allow'.

(8) Tara=ne Amu=ko hathi pinc **kar-va le-ne**
Tara=ERG Amu=DAT elephant.M.SG.NOM pinch do-CAUS take-INF.OBL
**di-ya**
give-PERF.M.SG
Tara let Amu have the elephant pinched (completely). (Butt et al. 2010: 1)

In Lowe's approach, the basic format of the meaning constructors of both the permissive and the causative predicates would be the same, and no mechanism is suggested to distinguish which one applies to the other in the semantics. Andrews (2018) proposes a solution to this problem which treats f-structures as hybrid objects. A hybrid object is an AVM that includes both an attribute-value pair and a set

of one or more AVMs (Dalrymple et al. 2019: 49). In a hybrid object, distributive features are those that must hold true of all AVMs in the set, as well as of the AVM that contains the set. Andrews creates a hybrid object by positing a phrase structure rule like the one in example (9) in which the f-structure of the main verb is placed in a singleton set in the f-structure of the light verb. This results in an f-structure like the one in Figure 10 in which the PRED of the light verb and main verb appear in different f-structures, yet those f-structure must share any distributive features.[3]

(9)
$$\begin{array}{lcll} \text{VP} & \rightarrow & \text{V} & \text{V} \\ & & \downarrow \in \uparrow & \uparrow = \downarrow \end{array}$$

(10) **F-structure of example (1) based on Andrews (2018)**

$$\begin{bmatrix} \text{PRED} & \text{‘let’} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{‘mother’} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{‘books’} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{‘child’} \end{bmatrix} \\ \left\{ \begin{bmatrix} \text{PRED} & \text{‘read’} \end{bmatrix} \right\} & \end{bmatrix}$$

This allows a named entity (e.g. %G) in the lexical entry to constrain which s-structure elements can be plugged into the meaning of the light verb, as in example (11). The light verb can only consume s-structure attributes that are projected by the f-structure which is in the named f-structure, the one inside the singleton set.

(11) $\%\text{G} \in \uparrow$
$\lambda P.\lambda y.\lambda x.\lambda e.let(x, y, P(y, e)) :$
$[(\%G_\sigma \ \text{ARG}1) \multimap (\%G_\sigma \ \text{EV}) \multimap \%G_\sigma] \multimap$
$(\uparrow_\sigma \ \text{ARG}3) \multimap (\uparrow_\sigma \ \text{ARG}1) \multimap (\uparrow_\sigma \ \text{EV}) \multimap \uparrow_\sigma$

Since Andrews’ (2018) proposal greatly expands the use of hybrid objects in f-structure, it requires a method for stipulating whether features are distributive or non-distributive on a construction-by-construction basis. Andrews (2018: 144) proposes a solution called “undersharing” but also states that “such undersharing specifications are theoretically somewhat undesirable.”

[3]The motivation for distributive and nondistributive features comes from analyses of feature resolution in coordination. For example, King and Dalrymple (2004) analyze the sentence *This boy and girl eat pizza.* Note that the determiner *this* as well as each of the nouns in the conjoined noun phrase must all be a singular form. Despite this, the verb form, *eat*, is not a third-person singular form, but a plural form. King and Dalrymple (2004) propose that there are two types of agreement. In the agreement internal to the noun phrase, the singular feature must hold true of each part of the phrase which contains a determiner and a set of conjoined nouns. It is a distributive feature. However, the plural feature that is reflected in the verb agreement does not have to hold of any of the elements of the phrase, only of the phrase itself. It is a nondistributive feature.

# 6 New proposal

## 6.1 F-structure

A similar, but simpler solution is to assume that the main verb's f-structure is embedded as the value of a grammatical function, arbitrarily labeled EP (cf. Lowe 2015: 422), as in example (12).[4] Since Completeness and Coherence are handled in glue semantics, a PRED can appear in an f-structure with no grammatical functions. Placing all grammatical functions in the f-structure of the light verb accounts for the emprical evidence showing that the f-structure only has a single set of grammatical functions (Section 2).

(12) **Proposed f-structure of example (1)**

$$\begin{bmatrix} \text{PRED} & \text{'let'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'mother'} \end{bmatrix} \\ \text{OBJ}_\theta & \begin{bmatrix} \text{PRED} & \text{'child'} \end{bmatrix} \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{'books'} \end{bmatrix} \\ \text{EP} & \begin{bmatrix} \text{PRED} & \text{'read'} \end{bmatrix} \end{bmatrix}$$

The embedded f-structure allows glue semantics to refer to the main verb by a local name. The meaning construction is identical to that in example (11), except that the local name is defined as %G = (↑ EP). Like the proposal of Andrews (2018), this allows the glue semantics to specify a particular structure as its input, but without having to adopt the additional complication of modeling the f-structure of complex predicates as a hybrid object. In the approach proposed here, no theoretical adjustments are needed other than those already proposed to adopt glue semantics as a replacement for Completeness and Coherence.

The proposed phrase structure rule for Urdu complex predicates is shown in example (13).[5] The node of the main verb is annotated as the value of EP of the verb that heads its constituent. Note that the EP node can also be a noun in the case of light verbs that take a nominal complement as in the causative verb in example (8).

(13) **Proposed PS rule for Urdu complex predicates**

$$\begin{array}{lcccc} \text{V} & \rightarrow & \{\text{V} \mid \text{N}\} & & \text{V} \\ & & (\uparrow \text{EP}) = \downarrow & & \uparrow = \downarrow \end{array}$$

[4] I leave open the question of what kind of f-structure attribute is needed for this function. In some languages, it may be possible to use the attribute COMP, but this depends on the particular analysis of each language, and any cross-linguistic claims made about the nature of particular f-structure attributes.

[5] The phrase structure rules in example (13) and elsewhere ignore constraints on the level of structure. Butt (1995) labels this mother node V′ although it is never dominated by a VP. A formal solution to how to model levels of structure that do not conform to the standard two levels proposed in X-bar theory can be found in minimal c-structure (Lowe and Lovestrand 2020).

The phrase structure rule in example (13) indicates that the light verb and main verb form a constituent in the c-structure. This analysis is confirmed by the fact that the light verb and main verb can "scramble" appearing in a non-sentence final position as long as they remain adjacent to each other, as in example (14).

(14) anjum=ne [**likʰ-ne** **d-ii**] ciṭṭʰii saddaf=ko
Anjum=ERG write-INF.OBL give-PRF.F.SG note(NOM) Saddaf=DAT
Anjum let Saddaf write a note. (Butt 1995: 46)

Butt (1995) shows that there are two types of c-structures for complex predicates in Urdu (see also Butt 1994). In addition to the main verb and light verb forming a verbal cluster constituent, it is also possible for the main verb and the object to form a constituent, as in example (15). This requires another kind of phrase structure rule, the one proposed in example (16).

(15) anjum=ne d-ii saddaf=ko [**ciṭṭʰii** **likʰ-ne**]
Anjum=ERG give-PRF.F.SG Saddaf=DAT note(NOM) write-INF.OBL
Anjum let Saddaf write a note. (Butt 1995: 46)

(16) **Proposed PS rule for Urdu complex predicates**

$$V \rightarrow \underset{(\uparrow \text{OBJ}) = \downarrow}{N} \quad \underset{(\uparrow \text{EP}) = \downarrow}{V}$$

Avery Andrews (personal communication) points out that the constituents licensed by the phrase structure rules in examples (13) and (16) cannot co-occur because both verbs would contribute a PRED attribute to the same f-structure, the value of EP. This predicts that in the case of very complex predicates in Urdu with more than one light verb, the light verbs must always be adjacent. It seems that this prediction will not bear out, however, more investigation is needed to determine what kind of constituent structures are permitted under what conditions (Miriam Butt, personal communication). This is a noteworthy gap in the empirical coverage of the approach to complex predicates proposed in this paper.

## 6.2 S-structure

The f-structure proposed in Section 6.1 allows a reasonably straightforward way to construct an s-structure of complex predicates that models the semantic composition of the construction. Asudeh and Giorgolo (2012) propose that the links between f-structure and s-structure be constrained by equations in the lexical entry of the predicate, as in example (17). Ideally, these constraints should be encoded in templates in the most generalized manner possible (Findlay 2016), but for immediate purposes the equations are only meant to represent relevant aspects of the lexical entry of the light verb in the permissive complex predicate.

(17) *dii* 'let'

a. (↑ PRED) = 'let'

b. ($\uparrow_\sigma$ REL) = let

c. ((↑ SUBJ)$_\sigma$ = ($\uparrow_\sigma$ ARG1))

d. ((↑ OBJ$_\theta$)$_\sigma$ = ($\uparrow_\sigma$ ARG3))

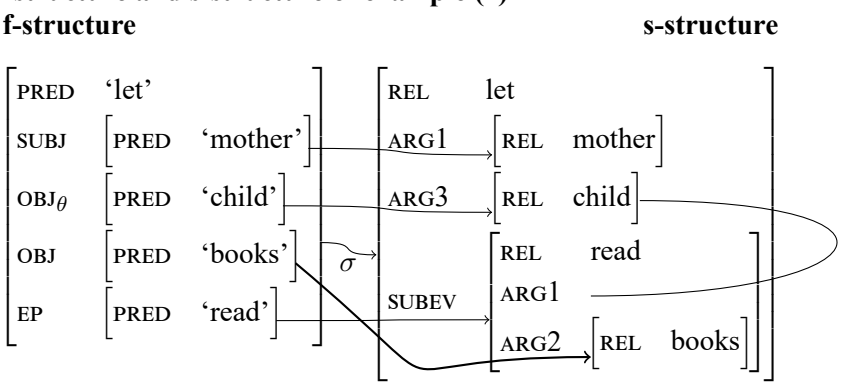
e. (↑ EP)$_\sigma$ = ($\uparrow_\sigma$ SUBEVENT )

f. ($\uparrow_\sigma$ ARG3) = ($\uparrow_\sigma$ SUBEVENT ARG1)

g. ((↑ OBJ)$_\sigma$ = ($\uparrow_\sigma$ SUBEVENT$^+$ ARG2))

The first two equations in example (17) state what the PRED and REL values are. The third equation links the SUBJ with ARG1 and the fourth equation links the OBJ$_\theta$ with ARG3. These are standard links that any general account of argument realization must include. The fifth equation embeds the semantic structure of the main verb, the s-structure projected from the value of EP, with the value of a SUBEVENT in s-structure.[6] The sixth equation creates a link in s-structure between the ARG3 and the ARG1 of the SUBEVENT—the permitee and the agent (as in example (3)). Note that the verb 'let' does not have an ARG2 in its s-structure, but the final equation links a grammatical function, OBJ, directly to an ARG2 in the value of a SUBEVENT. This linking equation is very similar to a standard equation needed in any theory of argument realization, with the difference being that one or more SUBEVENT attributes appear in the link between the grammatical function and the argument. Example (18) shows the f-structure and s-structure of example (1) under this analysis.

(18) **F-structure and s-structure of example (1)**

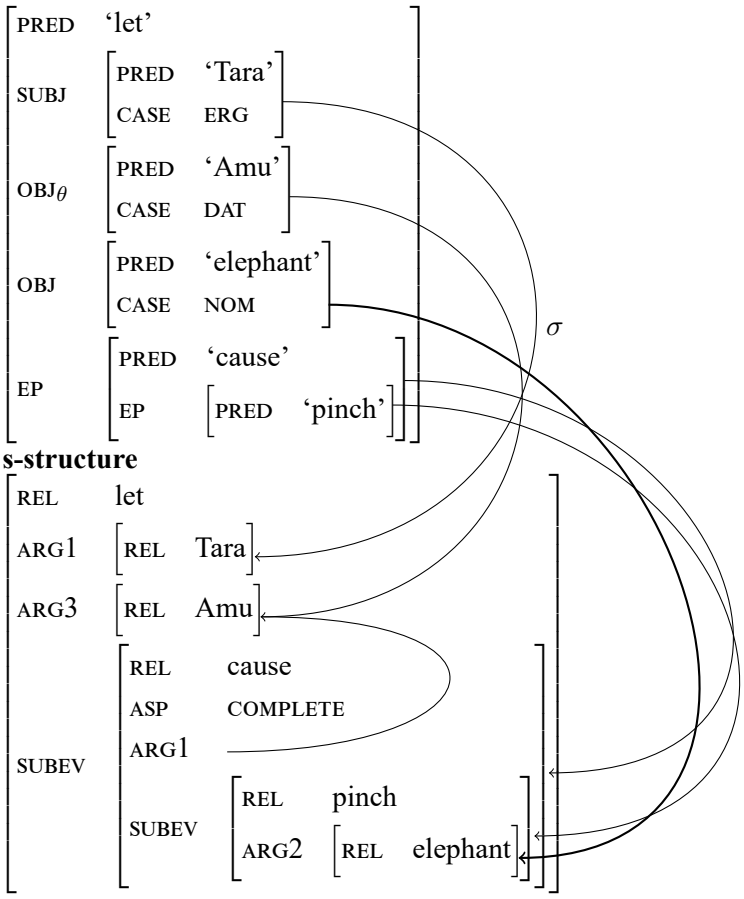



The lexical entry in example (17) also allows a representation of the s-structure of the very complex predicate in example (8) as shown in example (19). In this s-structure, the OBJ is linked to an ARG2 that is embedded in two SUBEVENTS. This flexibility in allowing different levels of embedding is allowed by the Kleene plus sign (one or more) in the equation in example (17g). The use of a regular expression

[6]The attribute SUBEVENT is a general label without any particular claims made about whether s-structure should only have one general type of attribute to embed other predicates, or whether there are several types.

to capture this kind of indeterminacy is not so different from the use of similar equations which place a Kleene star on a grammatical function to allow long distance dependencies in f-structure (Dalrymple et al. 2019: 207-208).

(19) **F-structure and s-structure of example (8)**




This analysis can be extended to cover another version of the Urdu permissive complex predicate, the "allow-to-happen" permissive. The permissive complex predicates discussed above (examples (1) and (8)) are the "allow-to-do" type in which the permission is directed towards a particular person or other animate argument. In constrast, in the "allow-to-happen" permissive, an event is allowed to take place (or not) without respect to permission being granted to any particular person, as in example (20).

(20) ḍakṭar=ne mɑriz=ko buxar a-ne nɑhĩ
doctor.SG=ERG patient.SG=DAT fever.M.SG.NOM come-INF.OBL not
di-ya
give-PRF.M.SG
'The doctor did not let the patient get a fever.' (Butt 2014: 22)

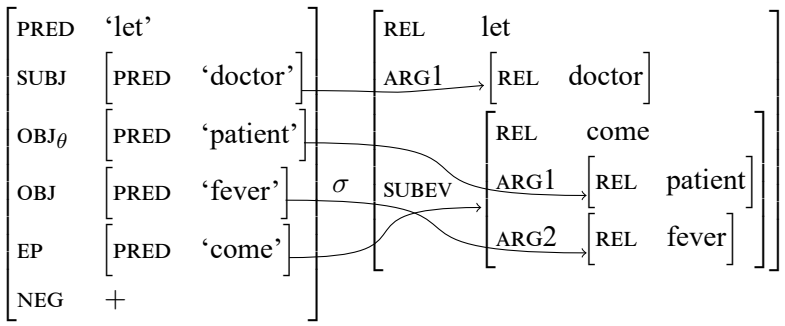
Butt's (2014: 22) analysis of the "allow-to-happen" permissive is to posit a modified lexical entry for the light verb *de* 'give' in this context. Along those lines, the lexical entry proposed for the permissive light verb in this paper can also be modified for the "allow-to-happen" interpretation. The assumption is that the difference in meaning reflects the absence of an ARG3 in the semantics. The first adjustment to the meaning constructor is that the *y* variable in example (21) only appears once in the semantics, as an argument of *P*, whereas above, in example (6), it also appears as an argument of the permissive predicate *let*. The effect is that the permissive predicate is treated as a two-place predicate, rather than a three-place predicate. The second adjustment is to remove the ARG3 from the glue expression, and replace it with an ARG1 embedded in a SUBEVENT.

(21) %G = (↑ EP)
$\lambda P.\lambda y.\lambda x.\lambda e.let(x, P(y, e))$ :
$[(\%G_\sigma$ ARG1$)\multimap(\%G_\sigma$ EV$)\multimap\%G_\sigma]\multimap$
$(\uparrow_\sigma$ SUBEVENT ARG1$)\multimap(\uparrow_\sigma$ ARG1$)\multimap(\uparrow_\sigma$ EV$)\multimap\uparrow_\sigma$

The equations linking f-structure and s-structure also need to be adjusted. The fourth equation in example (22) links the secondary object directly to the ARG1 of the subevent. This contrast with the lexical entry in example (17) where the $\text{OBJ}_\theta$ is linked to an ARG3 which is itself linked to the ARG1 of a SUBEVENT. The f-structure and s-structure resulting from this analysis are shown in example (23).

(22) *dii* 'let' ("allow-to-happen")
a. (↑ PRED) = 'let'
b. ($\uparrow_\sigma$ REL) = let
c. ((↑ SUBJ)$_\sigma$ = ($\uparrow_\sigma$ ARG1))
d. ((↑ $\text{OBJ}_\theta$)$_\sigma$ = ($\uparrow_\sigma$ SUBEVENT ARG1))
e. (↑ EP)$_\sigma$ = ($\uparrow_\sigma$ SUBEVENT )
f. ((↑ OBJ)$_\sigma$ = ($\uparrow_\sigma$ SUBEVENT$^+$ ARG2) )

(23) **F-structure and s-structure of example (20)**




An additional link is needed to account for complex predicates that express indirect causation, as in example (24). In this example the agent of the main predicate,

'pinch', can optionally be expressed in an instrumental phrase. Instead of an $\text{OBJ}_\theta$ linked to the ARG1 of the SUBEVENT as in example (22), an $\text{OBL}_{agent}$ is linked to that argument, as shown in example (25).

(24) ɑmu=ne (bɑcce=se) hat$^{h}$i pınc kar-va-ya
Amu=ERG child.OBL=INST elephant.M.SG.NOM pinch do-CAUS-PERF.M.SG
'Amu had the elephant pinched (by the child).' (Butt et al. 2010: 3)

(25) $((\uparrow_\sigma \text{OBL}_{agent}) = (\uparrow_\sigma \text{SUBEVENT}^{+} \text{ARG1}))$

# 7 XLE implementation

The analysis of Urdu permissive complex predicates proposed in this paper has been partially implemented in a mini-grammar in XLE. The major shortcoming of the current implementation is that it does not include a model of glue semantics. The result is that the parses allow many ambiguities which would be accounted for by the glue semantics. This shortcoming can likely be resolved by implementing a version of glue for XLE currently under development (Dalrymple et al. 2020). The space and format limitation of this paper prevent a detailed look at the XLE implementation, but f-structure and s-structure resulting from parsing example (8) in XLE is shown in Figure 2.

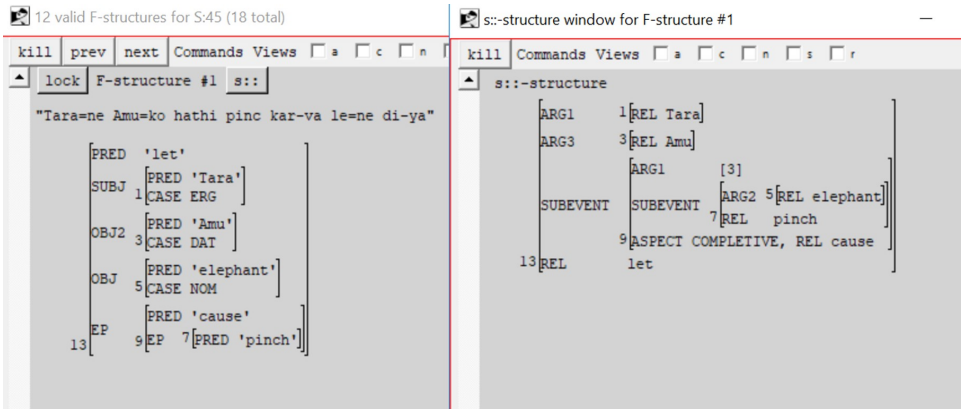



Figure 2: XLE parsing of example (8)

# 8 Conclusion

The work of Butt (1995) and others on complex predicates has made it clear that the original formulation of Lexical-Functional Grammar (Bresnan and Kaplan 1982) needs to be amended, but how? The most well-developed proposals on complex predicates in LFG have focused on removing the requirement that semantic forms be non-unifiable, allowing them to unify in the syntax. The alternative is to remove

the Completeness and Coherence constraints, and allow glue semantics to handle constraints on what arguments appear, giving much more flexibility in f-structure representations. Although this was first proposed by Dalrymple et al. (1993), only in more recent years has the proposal been more fully developed (Asudeh and Giorgolo 2012, Lowe 2015, 2019, Andrews 2018). This paper adds to the development of this approach to complex predicates. It has now been established that a glue-based approach allows a model of complex predicates in which semantic forms remain non-unifiable. The model captures the empirical facts related to f-structure, and allows s-structure to not only model argument structure, but also to be the locus of a more complete representation of syntactically relevant semantic information. It also seems that this approach can be implemented in XLE, pending further development of incorporating glue meaning constructors in XLE.

However, the approach also makes a prediction that the components of a very complex predicate (two or more light verbs) will necessarily appear in a cluster in the c-structure. This prediction does not seem to hold up against the facts of Urdu (Miriam Butt, personal communication). Nonetheless, the approach can be further tested against other types of complex predicates in Urdu and other languages, such as the Romance complex predicates analyzed by Andrews (2018), potentially revealing other gaps or new insights that could resolve the apparent issue.

Another weakness of this approach is that it remains relatively stipulative in regards to the lexical entries of light verbs. More cross-linguistic work is needed to make any generalizations about how complex predicates fit into the templatic approach to representing links between f-structure and s-structure (Findlay 2016, 2020). More generally speaking, the connected s-structure used in this approach is relatively undeveloped. The potential of a connected s-structure for representing lexical semantics and argument realization remains unexplored. For example, can s-structure be used to model the semantic features that Butt (1995) uses Lexical Conceptual Structures (Jackendoff 1990) to model?

# References


Alsina, Alex. 1993. *Predicate composition: A theory of syntactic function alternations*. PhD dissertation, Stanford University.

Alsina, Alex. 1997. A theory of complex predicates: Evidence from causatives in Bantu and Romance. In Alex Alsina, Joan Bresnan and Peter Sells (eds.), *Complex predicates*, pages 203–246, Stanford: CSLI Publications.

Andrews, Avery D. 2008. The role of PRED in LFG+glue. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, pages 47–67, CSLI Publications.

Andrews, Avery D. 2018. Sets, heads and spreading in LFG. *Journal of Language Modelling* 6(1), 131–174.

Asudeh, Ash and Giorgolo, Gianluca. 2012. Flexible composition for optional and derived arguments. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG12 Conference*, pages 64–84, CSLI Publications.

Asudeh, Ash, Giorgolo, Gianluca and Toivonen, Ida. 2014. Meaning and valency. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG14 Conference*, pages 68–88, CSLI Publications.

Bresnan, Joan and Kanerva, Jonni M. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* pages 1–50.

Bresnan, Joan and Kaplan, Ronald M. 1982. Introduction: Grammars as Mental Represenations of Language. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, pages xvii–lii, The MIT Press.

Butt, Miriam. 1994. Complex predicate scrambling in Urdu. In Miriam Butt, Tracy Holloway King and Gillian Ramchand (eds.), *Theoretical Perspective on Word Order in South Asian Languages*, pages 67–90, Stanford, CA: CSLI Publications.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Stanford, CA: CSLI Publications.

Butt, Miriam. 1997. Complex predicates in Urdu. In Alex Alsina, Joan Bresnan and Peter Sells (eds.), *Complex predicates*, Stanford, CA: CSLI Publications.

Butt, Miriam. 2014. Control vs. complex predication: Identifying non-finite complements. *Natural Language & Linguistic Theory* 32(1), 165–190.

Butt, Miriam, Dalrymple, Mary and Frank, Anette. 1997. An architecture for linking theory in LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*, pages 1–16.

Butt, Miriam and King, Tracy Holloway. 2006. Restriction for Morphological Valency Alternations : The Urdu Causative. In Miriam Butt, Mary Dalrymple and Tracy Holloway King (eds.), *Intelligent Linguistic Architectures: Variations on Themes by Ronald M. Kaplan*, pages 235–258, Stanford, CA: CSLI Publications.

Butt, Miriam, King, Tracy Holloway and Maxwell III, John T. 2003. Complex predicates via restriction. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG03 Conference*, pages 92–104, CSLI Publications.

Butt, Miriam, King, Tracy Holloway and Ramchand, Gillian. 2010. Complex predication: How did the child pinch the elephant. In Linda Uyechi and Lian-Hee Wee (eds.), *Reality Exploration and Discovery: Pattern Interaction in Language & Life. A Festschrift for KP Mohanan*, pages 231–256, Stanford, CA: CSLI Publications.

Butt, Miriam and Ramchand, Gillian. 2005. Complex aspectual structure in Hindi/Urdu. In Nomi Erteschik-Shir and Tova Rapoport (eds.), *The syntax of aspect: Deriving thematic and aspectual information*, pages 117–153, Oxford: Oxford University Press.

Dalrymple, Mary, Hinrichs, Angie, Lamping, John and Saraswat, Vijay. 1993. The resource logic of complex predicate interpretation. In *Proceedings of the 1993 Republic of China Computational Linguistics Conference (ROCLING)*, pages 3–21, Computational Linguistics Society of Republic of China, Hsitou National Park, Taiwan. Also published as Xerox Technical Report ISTL-NLTT-1993-08-03.

Dalrymple, Mary, Lowe, John and Mycock, Louise. 2019. *The Oxford Reference Guide to Lexical Functional Grammar*. Oxford: Oxford University Press.

Dalrymple, Mary, Patejuk, Agnieszka and Zymla, Mark-Matthias. 2020. XLE+Glue – A new tool for integrating semantic analysis in XLE. In Miriam Butt and Ida Toivonen (eds.), *Proceedings of the LFG20 Conference*, Stanford, CA: CSLI Publications.

Findlay, Jamie. 2016. Mapping theory without argument structure. *Journal of Language Modelling* 4(2), 293–338.

Findlay, Jamie Y. 2020. Lexical Mapping Theory and the anatomy of a lexical entry. In Miriam Butt and Ida Toivonen (eds.), *Proceedings of the LFG20 Conference*, Stanford, CA: CSLI Publications.

Jackendoff, Ray. 1990. *Semantic structures*. MIT press.

King, Tracy Holloway and Dalrymple, Mary. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40(1), 69–104.

Kuhn, Jonas. 2001. Resource sensitivity in the syntax-semantics interface and the German split NP construction. In W. Detmar Meurers and Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax*, pages 75–117, Stanford, CA: CSLI Publications.

Lovestrand, Joseph. 2018. *Serial verb constructions in Barayin: Typology, description and Lexical-Functional Grammar*. PhD dissertation, University of Oxford.

Lowe, John J. 2015. Complex predicates: An LFG+glue analysis. *Journal of Language Modelling* 3(2), 413–462.

Lowe, John J. 2019. Argument alternations in complex predicates: an LFG+ glue perspective. *Linguistic Issues in Language Technology* 17(2).

Lowe, John J. and Lovestrand, Joseph. 2020. Minimal phrase structure: A new formalized theory of phrase structure. *Journal of Language Modelling* 8(1).

Przepiórkowski, Adam. 2017. Hierarchical lexicon and the argument/adjunct distinction. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG17 Conference*, pages 348–367, Stanford, CA: CSLI Publications.